

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model
Run on: October 22, 2004, 13:36:34 ; Search time 158 seconds
(without alignments)
95.358 Million cell updates/sec

Title: US-09-763-982B-1
Perfect score: 243
Sequence: 1 NQSNFGFMKGNFGRRSG.....GGGQYFAKPRNQGGYGGC 42

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	243	100.0	42	3	AY82221	Modified	AY82221 Modified
2	228	93.8	76	5	AB24273	Human ova	Abp24273 Human ova
3	228	93.8	250	4	ABG03369	Novel hum	Abg03369 Novel hum
4	228	93.8	296	4	ABG09555	Novel hum	Abg09555 Novel hum
5	228	93.8	296	4	ABG04261	Novel hum	Abg04261 Novel hum
6	228	93.8	296	4	ABG03366	Novel hum	Abg03366 Novel hum
7	228	93.8	296	4	ABG15176	Novel hum	Abg15176 Novel hum
8	228	93.8	296	4	AU29983	Novel hum	Au29983 Novel hum
9	228	93.8	298	7	ADJ69744	Human hea	Adj69744 Human hea
10	228	93.8	319	7	AD57408	Rat Prote	Ad57408 Rat Prote
11	228	93.8	320	2	AAW55828	Human het	Aaw55828 Human het
12	228	93.8	320	4	AB48966	Human het	Ab48966 Human het
13	228	93.8	320	4	AB381934	Marmosat	Ab381934 Marmosat
14	228	93.8	320	4	ABG00955	Novel hum	Abg00955 Novel hum
15	228	93.8	320	5	AB57241	Mouse isc	Ab57241 Mouse isc
16	228	93.8	320	8	AD117083	Human hnr	Ad117083 Human hnr
17	228	93.8	351	7	AD60577	Human con	Ad60577 Human con
18	228	93.8	369	4	ABG15179	Novel hum	Abg15179 Novel hum
19	228	93.8	372	6	AB052967	Human spl	Ab052967 Human spl
20	228	93.8	470	4	ABG04264	Novel hum	Abg04264 Novel hum
21	228	93.8	474	4	ABG04458	Novel hum	Abg04458 Novel hum
22	228	93.8	1205	4	ABG15312	Novel hum	Abg15312 Novel hum
23	228	93.8	1205	4	ABG05068	Novel hum	Abg05068 Novel hum
24	225	92.6	95	4	ABG15177	Novel hum	Abg15177 Novel hum
25	225	92.6	95	4	ABG27119	Novel hum	Abg27119 Novel hum

26	225	92.6	95	4	ABG09556	Novel hum	Abg09556 Novel hum
27	225	92.6	312	4	ABG06528	Novel hum	Abg06528 Novel hum
28	225	92.6	320	7	AD57410	Human Pro	Ad57410 Human Pro
29	222	91.4	1214	4	ABG27121	Novel hum	Abg27121 Novel hum
30	218	89.7	324	6	ABU11521	Human MDD	Abu11521 Human MDD
31	216	88.9	38	3	AAV67200	Novel hum	Aav67200 Novel hum
32	216	88.9	38	3	AAV82223	M9 nuclea	Aav82223 M9 nuclea
33	216	88.9	38	4	AAE02957	Human RNP	Aae02957 Human RNP
34	216	88.9	44	2	AAW08405	M9 includ	Aaw08405 M9 includ
35	212.5	87.4	41	6	ABP56848	Peptide M	Abp56848 Peptide M
36	197	81.1	254	4	ABG16296	Novel hum	Abg16296 Novel hum
37	197	81.1	254	4	ABG04260	Novel hum	Abg04260 Novel hum
38	197	81.1	254	4	ABG04455	Novel hum	Abg04455 Novel hum
39	197	81.1	237	4	ABG00954	Novel hum	Abg00954 Novel hum
40	193	79.4	558	4	ABG15178	Novel hum	Abg15178 Novel hum
41	193	79.4	558	4	ABG09557	Novel hum	Abg09557 Novel hum
42	193	79.4	558	4	ABG04262	Novel hum	Abg04262 Novel hum
43	184	75.7	359	4	ABG04454	Novel hum	Abg04454 Novel hum
44	184	75.7	370	4	ABG05066	Novel hum	Abg05066 Novel hum
45	179	73.7	133	4	ABG03368	Novel hum	Abg03368 Novel hum

ALIGNMENTS

RESULT 1
AAV82221
ID AAV82221 standard; peptide; 42 AA.
XX AC AAV82221;
XX XX
DT 13-JUN-2000 (first entry)
DE Modified M9 nuclear targeting peptide SEQ ID NO:1.
KW Human; M9; heteronuclear ribonuclear protein type 1; hnRNP A1; NTP;
KW nuclear targeting peptide; peptide scaffold; gene transfer; NLS;
KW nuclear localisation signal; antiarteriosclerotic; vasotropic;
KW Gene therapy; atherosclerosis; restenosis; angioplasty.
XX Homo sapiens.
XX WO200012114-A1.
XX PD 09-MAR-2000.
XX PF 01-SEP-1999; 99WO-US020122.
XX PR 01-SEP-1998; 98US-0098791P.
XX (UYPE-) UNIV PENNSYLVANIA.
XX Diamond SL;
XX WPI; 2000-256490/22.

Delivering molecules to nuclei of eukaryotic cells for use in gene transfer methods uses a nuclear targeting peptide which contains a non-classical nuclear localization signal.
Claim 8; Page 17; 40pp; English.
The present invention describes a composition for enhancing delivery of a molecule to the nucleus of eukaryotic cells, comprising a nuclear targeting peptide (NTP) containing a non-classical nuclear localisation signal (NLS). The composition can have antiarteriosclerotic and vasotropic activities, and can be used in gene therapy. The composition is used to treat a patient having a condition associated with lack of expression of a selected nucleic acid sequence. The compositions are particularly useful for arterial gene transfer, to treat atherosclerosis and restenosis following angioplasty. The present sequence represents a specifically claimed NTP which comprises the human heteronuclear ribonuclear protein type 1 (hnRNP A1) M9 epitope with a carboxy terminal

CC cysteine residue
 XX Sequence 42 AA;
 SQ

Query Match 100.0%; Score 243; DB 3; Length 42;
 Best Local Similarity 100.0%; Pred. No. 6.5e-21;
 Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NQSNFGPMKGNFGGRSSGPGYGGGQYFAKPRNQGGYGGC 42
 DB 1 NQSNFGPMKGNFGGRSSGPGYGGGQYFAKPRNQGGYGGC 42

RESULT 2
 ABP42473
 ID ABP42473 standard; protein; 76 AA.
 AC ABP42473;
 DT 22-AUG-2002 (first entry)
 XX Human ovarian antigen HOCQM24, SEQ ID NO:3605.
 DE Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
 KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
 KW inflammatory condition; immune disorder; blood disorder;
 KW cardiovascular disorder; respiratory disorder; neurological disorder;
 KW gastrointestinal disorder; urinary system disorder; drug screening;
 KW gene therapy; chromosome mapping; forensic analysis;
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
 KW antiinflammatory; gynaecological; reproductive.
 XX Homo sapiens.
 OS
 XX WO200200677-A1.
 PN
 XX 03-JAN-2002.
 PD
 XX 07-JUN-2001; 2001WO-US018569.
 PF
 XX 07-JUN-2000; 2000US-0209467P.
 PR
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 PI Birse CE, Rosen CA;
 XX WPI; 2002-147878/19.
 DR
 XX N-PSDB; ABQ55550.
 XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,
 PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian
 PT cancer), immune disorders, cardiovascular disorders and neurological
 PT diseases.
 XX
 PS Claim 11; SEQ ID NO 3605; 2922pp; English.
 XX
 XX The invention relates to 2175 novel human ovarian antigens (ABP41054-
 CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical
 CC to the sequences of the invention. The invention additionally relates to
 CC recombinant vectors and host cells comprising human ovarian antigen
 CC polynucleotides, antibodies against human ovarian antigens, and the use
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
 CC treating, prognosing or preventing various ovary and/or breast-related
 CC disorders. Such conditions include ovarian cancer and breast cancer, and
 CC metastatic tumours of ovarian or breast origin, reproductive system
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
 CC vaginitis), immune disorders (e.g., congenital and acquired

CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
 CC respiratory disorders, neurological disorders, gastrointestinal disorders
 CC and urinary system disorders. Ovarian antigen polypeptides and
 CC polynucleotides may also be used in screening for compounds which
 CC modulate ovarian antigen expression or activity. The polynucleotides may
 CC further be used for gene therapy, chromosome mapping, in the
 CC identification of individuals and in forensic analysis, and the
 CC polypeptides may be used as food additives or to prepare antibodies
 CC useful in disease diagnosis, drug targeting and phenotyping. The present
 CC sequence represents a human ovarian antigen of the invention. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 76 AA;
 Query Match 93.8%; Score 228; DB 5; Length 76;
 Best Local Similarity 100.0%; Pred. No. 6.4e-19;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NQSNFGPMKGNFGGRSSGPGYGGGQYFAKPRNQGGYGG 40
 DB 24 NQSNFGPMKGNFGGRSSGPGYGGGQYFAKPRNQGGYGG 63

RESULT 3
 ABG03369
 ID ABG03369 standard; protein; 250 AA.
 XX ABG03369;
 AC ABG03369;
 DT 13-FEB-2002 (first entry)
 XX Novel human diagnostic protein #3360.
 DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX Homo sapiens.
 OS
 XX WO200175067-A2.
 PN
 XX 11-OCT-2001.
 PD
 XX 30-MAR-2001; 2001WO-US008631.
 PF
 XX 31-MAR-2000; 2000US-00540217.
 PR
 XX 23-AUG-2000; 2000US-00649167.
 XX (HYSE-) HYSEQ INC.
 PA
 PI Drmanac RT, Liu C, Tang YT;
 XX WPI; 2001-639362/73.
 DR
 XX N-PSDB; AAS67556.
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 20; SEQ ID NO 33728; 103pp; English.
 XX
 XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food

CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 250 AA;

Query Match 93.8%; Score 228; DB 4; Length 250;
 Best Local Similarity 100.0%; Pred. No. 2.1e-18;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NQSSNFGPMKGNFGRRSGPYGGGGQYFAKPRNQGGYGG 40
 |||||
 DB 198 NQSSNFGPMKGNFGRRSGPYGGGGQYFAKPRNQGGYGG 237

RESULT 4

ABG09555
 ID ABG09555 standard; protein; 296 AA.

XX AC ABG09555;

XX DT 13-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #9546.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS73742.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.

XX PS Claim 20; SEQ ID NO 39914; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders

CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 296 AA;

Query Match 93.8%; Score 228; DB 4; Length 296;
 Best Local Similarity 100.0%; Pred. No. 2.5e-18;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NQSSNFGPMKGNFGRRSGPYGGGGQYFAKPRNQGGYGG 40
 |||||
 DB 244 NQSSNFGPMKGNFGRRSGPYGGGGQYFAKPRNQGGYGG 283

RESULT 5

ABG04261
 ID ABG04261 standard; protein; 296 AA.

XX AC ABG04261;

XX DT 13-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #4252.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS68448.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.

XX PS Claim 20; SEQ ID NO 34620; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 296 AA;

Query Match 93.8%; Score 228; DB 4; Length 296;
 Best Local Similarity 100.0%; Pred. NO. 2.5e-18;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NQSNFGPMKGNFGGRSSGPGYGGGQYFAKPRNQGGYGG 40
 |||||
 Db 244 NQSNFGPMKGNFGGRSSGPGYGGGQYFAKPRNQGGYGG 283

RESULT 6
 ABG03366
 ID ABG03366 standard; protein; 296 AA.

XX AC ABG03366;

XX DT 13-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #337.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 XX food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS67553.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.

XX PS Claim 20; SEQ ID NO 33725; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC and to produce other types of data and products dependent on DNA and
 CC responsible for genetic disorders or other traits to assess biodiversity

CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 296 AA;

Query Match 93.8%; Score 228; DB 4; Length 296;
 Best Local Similarity 100.0%; Pred. NO. 2.5e-18;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NQSNFGPMKGNFGGRSSGPGYGGGQYFAKPRNQGGYGG 40
 |||||
 Db 244 NQSNFGPMKGNFGGRSSGPGYGGGQYFAKPRNQGGYGG 283

RESULT 7
 ABG15176
 ID ABG15176 standard; protein; 296 AA.

XX AC ABG15176;

XX DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #15167.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 XX food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS79363.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.

XX PS Claim 20; SEQ ID NO 45535; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic

CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 296 AA;

Query Match 93.8%; Score 228; DB 4; Length 296;
 Best Local Similarity 100.0%; Pred. No. 2.5e-18;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NOSSNFGPMKGNFGRRSGPYGGGQYFAKPRNQGGYGG 40
 |||||
 Db 244 NOSSNFGPMKGNFGRRSGPYGGGQYFAKPRNQGGYGG 283
 |||||

RESULT 8
 AAU29983
 ID AAU29983 standard; protein; 296 AA.
 AC AAU29983;
 DT 18-DEC-2001 (first entry)
 DE Novel human secreted protein #474.
 XY Human; vaccination; gene therapy; nutritional supplement;
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
 XX Homo sapiens.
 OS Homo sapiens.
 FN WO200179449-A2.
 PD 25-OCT-2001.
 XX 16-APR-2001; 2001WO-US008656.
 XX 18-APR-2000; 2000US-0052929.
 PR 26-JAN-2001; 2001US-00770160.
 XX (HYSE-) HYSEQ INC.
 PA Tang YT, Liu C, Drmanac RT;
 PI WPI; 2001-611725/70.
 DR Nucleic acids encoding a range of human polypeptides, useful in genetic
 PT vaccination, testing and therapy.
 XX Claim 20; Page 217; 765pp; English.

CC The invention relates to novel human secreted polypeptides. The
 CC polypeptides and antibodies to the polypeptides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of polypeptide. The polypeptides are also useful for
 CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC expressing the proteins are useful for identifying a therapeutic agent
 CC for use in treatment of a pathology related to aberrant expression or
 CC physiological interactions of the polypeptide. Vectors comprising the
 CC nucleic acids encoding the polypeptides and cells genetically engineered
 CC to express them are also useful for producing the proteins. The proteins
 CC are useful in genetic vaccination, testing and therapy, and can be used
 CC as nutritional supplements. They may be used to increase stem cell
 CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
 CC and/or nerve tissue growth or regeneration; immune suppression and/or
 CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
 CC AAU29510-AAU33304 represent the amino acid sequences of novel human
 CC secreted proteins of the invention
 XX
 SQ Sequence 296 AA;

Query Match 93.8%; Score 228; DB 4; Length 296;

Best Local Similarity 100.0%; Pred. No. 2.5e-18;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NOSSNFGPMKGNFGRRSGPYGGGQYFAKPRNQGGYGG 40
 |||||
 Db 244 NOSSNFGPMKGNFGRRSGPYGGGQYFAKPRNQGGYGG 283
 |||||

RESULT 9
 ADJ69744
 ID ADJ69744 standard; protein; 298 AA.
 AC ADJ69744;
 DT 06-MAY-2004 (first entry)
 XX Human heat mitochondrial protein as a therapeutic target SeqID1550.
 DE Mitochondrial; human; screening assay; diabetes mellitus;
 XX Huntington's disease; osteoarthritis; LHON;
 KW Leber's hereditary optic neuropathy; lactic acidosis and stroke; MELAS;
 KW mitochondrial encephalopathy ragged red fibre syndrome; MERRF; cancer;
 KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;
 KW neuroprotective; nontropic; antidiabetic; anticonvulsant; antiarthritic;
 KW osteopathic; ophthalmological; cytostatic.

XX Homo sapiens.
 OS Homo sapiens.
 XX WO2003087768-A2.
 FN 23-OCT-2003.
 PD 04-APR-2003; 2003WO-US010870.
 PF 12-APR-2002; 2002US-0372843P.
 XX 17-JUN-2002; 2002US-0389987P.
 PR 20-SEP-2002; 2002US-0412418P.
 XX (MITO-) MITOKOR.
 PA (BUCK-) BUCK INST AGE RES.
 XX Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;
 PI Warnock DE;
 PI WPI; 2003-845369/78.
 DR Identifying a mitochondrial target for drug screening assays and for
 XX treating diseases associated with altered mitochondrial function.
 PT comprises detecting a modified polypeptide in a sample and correlating
 PT with the disease.

XX Claim 1; SEQ ID NO 1550; 180pp; English.
 PS This invention relates to novel mitochondrial targets that can be used
 CC for therapeutic intervention in treating a disease associated with
 CC altered mitochondrial function. Specifically, it refers to a method for
 CC identifying proteins of the human heart mitochondrial proteome that are
 CC useful for drug screening assays, as well as therapeutic targets. The
 CC present invention describes a method for identifying such proteins that
 CC can be used in the treatment of various diseases associated with altered
 CC mitochondrial function including diabetes mellitus, Huntington's disease,
 CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
 CC encephalopathy, lactic acidosis and stroke (MELAS), myoclonic epilepsy
 CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
 CC compositions have neuroprotective, nontropic, antidiabetic,
 CC anticonvulsant, antiarthritic, osteopathic, ophthalmological and
 CC cytostatic activities. This polypeptide sequence is a human heart
 CC mitochondrial protein of the invention.
 XX
 SQ Sequence 298 AA;

Query Match 93.8%; Score 228; DB 7; Length 298;
 Best Local Similarity 100.0%; Pred. No. 2.5e-18;

Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NQSNFGPMKGNFGGRSSGPGYGGGQYFAKPRNQGGYGG 40
 DB 246 NQSNFGPMKGNFGGRSSGPGYGGGQYFAKPRNQGGYGG 285

RESULT 10
 ADE57408
 ID ADE57408 standard; protein; 319 AA.
 XX AC ADE57408;
 XX DT 29-JAN-2004 (first entry)
 XX DE Rat Protein P04256, SEQ ID NO 3269.
 XX KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
 XX KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
 XX OS Rattus norvegicus.
 XX FN WO2003016475-A2.
 XX PD 27-FEB-2003.
 XX PF 14-AUG-2002; 2002WO-US025765.
 XX PR 14-AUG-2001; 2001US-0312147P.
 XX PR 01-NOV-2001; 2001US-0346382P.
 XX PR 26-NOV-2001; 2001US-0333347P.
 XX PA (GEHO) GEN HOSPITAL CORP.
 XX PA (FARB) BAYER AG.
 XX PI Woolf C, D'urso D, Befort K, Costigan M;
 XX WPI; 2003-266312/26.
 XX DR GENBANK; P04256.
 XX PT New composition comprising two or more isolated polypeptides, useful for
 PT preparing a medicament for treating pain in an animal.
 XX PS Claim 1; Page; 1017pp; English.

The invention discloses a composition comprising two or more isolated rat
 or human polynucleotides or a polynucleotide which represents a fragment,
 derivative or allelic variation of the nucleic acid sequence. Also
 claimed are a vector comprising the novel polynucleotide, a host cell
 comprising the vector, a method for identifying a nucleotide sequence
 which is differentially regulated in an animal subjected to pain and a
 kit to perform the method, an array, a method for identifying an agent
 that increases or decreases the expression of the polynucleotide sequence
 that is differentially expressed in neuronal tissue of a first animal
 subjected to pain, a method for identifying a compound which regulates
 the expression of a polynucleotide sequence which is differentially
 expressed in an animal subjected to pain, a method for identifying a
 compound that regulates the activity of one or more of the
 polynucleotides, a method for producing a pharmaceutical composition, a
 method for identifying a compound or small molecule that regulates the
 activity in an animal of one or more of the polypeptides given in the
 specification, a method for identifying a compound useful in treating
 pain and a pharmaceutical composition comprising the one or more
 polypeptides or their antibodies. The polynucleotide or the compound that
 modulates its activity is useful for preparing a medicament for treating
 pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 therapy). The sequence presented is a rat protein (shown in Table 2 of
 the specification) which is differentially expressed during pain. Note:
 The sequence data for this patent did not form part of the printed
 specification, but was obtained in electronic form directly from WIPO at
 ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 319 AA;
 Query Match 93.8%; Score 228; DB 7; Length 319;
 Best Local Similarity 100.0%; Pred. No. 2.7e-18;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NQSNFGPMKGNFGGRSSGPGYGGGQYFAKPRNQGGYGG 40
 DB 267 NQSNFGPMKGNFGGRSSGPGYGGGQYFAKPRNQGGYGG 306

RESULT 11
 AAW55828
 ID AAW55828 standard; protein; 320 AA.
 XX AC AAW55828;
 XX DT 20-JUL-1998 (first entry)
 XX DE Human heterogeneous nuclear ribonucleoprotein core protein A1.
 XX KW Human; heterogeneous nuclear ribonucleoprotein; core protein; hnRNP;
 XX KW RNA binding protein; detection; coding region.
 XX OS Homo sapiens.
 XX PN JP10023893-A.
 XX PD 27-JAN-1998.
 XX PF 09-JUL-1996; 96JP-00179521.
 XX PR 09-JUL-1996; 96JP-00179521.
 XX PA (HITA) HITACHI LTD.
 XX WPI; 1998-152795/14.
 XX DR N-PSDB; AAV25978.
 XX PT Detection of protein code region on DNA base sequence - using method
 PT which reduces interference from noncoding region.
 XX PS Disclosure; Fig 2-5; 16pp; Japanese.

The present sequence represents the human heterogeneous nuclear
 ribonucleoprotein (hnRNP) core protein A1, which is used to exemplify the
 method of the present invention. The hnRNP A1 gene can be found on the
 genbank database accession number X12671, NID G32344. The present
 invention describes a method for the detection of a protein coding region
 in a DNA base sequence by judging the probability where a specific
 sequence will be found. The method comprises utilizing the appearance
 frequency in the coding and the noncoding regions. The method allows
 noise in the noncoding region to be reduced

SQ Sequence 320 AA;
 Query Match 93.8%; Score 228; DB 2; Length 320;
 Best Local Similarity 100.0%; Pred. No. 2.7e-18;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NQSNFGPMKGNFGGRSSGPGYGGGQYFAKPRNQGGYGG 40
 DB 268 NQSNFGPMKGNFGGRSSGPGYGGGQYFAKPRNQGGYGG 307

RESULT 12
 AAB48966
 ID AAB48966 standard; protein; 320 AA.
 XX AC AAB48966;
 XX DT 27-MAR-2001 (first entry)
 XX

DE Human heterogeneous nuclear ribonucleoprotein A1 (hnRNP A1).
 XX
 KW Human hnRNP A1; heterogeneous nuclear ribonucleoprotein A1;
 KW heterogeneous nuclear ribonucleoprotein core protein A1; p40CRS;
 KW mRNA processing; transport; stabilisation; alternative splicing;
 KW donor splice site selection; telomere biogenesis; oncogenesis;
 KW apoptosis-associated protein; cancer; tumour formation;
 KW expression inhibition; antisense therapy.
 XX
 OS Homo sapiens.
 XX
 PN US6165789-A.
 XX
 PD 26-DEC-2000.
 XX
 PF 27-OCT-1999; 99US-00428696.
 XX
 PR 27-OCT-1999; 99US-00428696.
 XX
 PA (ISIS-) ISIS PHARM INC.
 XX
 PI Monia BP, Cowsett LM;
 XX
 DR WPI; 2001-090484/10.
 DR N-PSDB; AAC92731.
 XX
 PT Novel antisense compound targeted to human hnRNP A1 which specifically
 PT hybridizes with and inhibits the expression of human hnRNP A1, useful for
 PT modulating the expression of hnRNP A1 in cells.
 XX
 PS Example 16; Col 43-48; 38pp; English.
 XX
 CC This sequence represents human heterogeneous nuclear ribonucleoprotein A1
 CC (hnRNP A1). hnRNP A1 (also known as heterogeneous nuclear
 CC ribonucleoprotein core protein A1 and p40CRS) is thought to function in
 CC the stabilisation, transport and processing (including alternative
 CC splicing) of newly synthesised mRNAs. It facilitates the annealing of
 CC single-stranded nucleic acids, modulates the binding of snRNPs to RNA
 CC intron sequences, and shuttles continuously between the nucleus and the
 CC cytoplasm acting as a carrier protein for mRNAs. hnRNP A1 also
 CC participates in telomere biogenesis, with low levels of hnRNP correlating
 CC with shortened telomeres. In addition, hnRNP A1 has also been classified
 CC as an apoptosis-associated protein on the basis that it is specifically
 CC cleaved into three fragments during antibody-mediated apoptosis. Due to
 CC its ability to control splicing events, particularly donor splice site
 CC selection, hnRNP A1 is implicated in the process of oncogenesis. The
 CC invention relates to antisense oligonucleotides targeted to the hnRNP A1
 CC gene, which inhibit its expression. A series of oligonucleotides
 CC (AAC92738-C92817) were designed to target different regions of the human
 CC hnRNP A1 mRNA, and were analysed for their effect on hnRNP A1 mRNA levels
 CC by quantitative real-time PCR. The oligonucleotides of the invention are
 CC useful for diagnosis, prevention and treatment of conditions associated
 CC with hnRNP A1 expression, such as cancer
 XX
 SQ Sequence 320 AA;
 Query Match 93.8%; Score 228; DB 4; Length 320;
 Best Local Similarity 100.0%; Pred. No. 2.7e-18;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NOSSNFGPMKGNFGRRSSGPGGGGQYFAKPRNQGGYGG 40
 DB 268 NOSSNFGPMKGNFGRRSSGPGGGGQYFAKPRNQGGYGG 307
 RESULT 13
 AAB81934
 ID AAB81934 standard; protein; 320 AA.
 XX
 AC AAB81934;
 XX
 DT 25-JUN-2001 (first entry)
 XX
 PF 25-JUN-2001 (first entry)
 XX

DE Marmoset vitamin D response element binding protein #1.
 XX
 KW Marmoset; vitamin D response element binding protein; VDRE-BP; lymphoma;
 KW vitamin D resistance; osteoporosis; hypercalcaemia; vitamin D toxicity;
 KW glucocorticoid mediated disorder; granuloma forming disease;
 KW vitamin D intoxication; steroid hormone hypersecretion; gene therapy.
 XX
 OS Saginus oedipus.
 XX
 PN WO200121649-A2.
 XX
 PD 29-MAR-2001.
 XX
 PF 20-SEP-2000; 2000WO-US025844.
 XX
 PR 22-SEP-1999; 99US-00400967.
 XX
 PA (CEDA-) CEDARS SINAI MEDICAL CENT.
 XX
 PI Adams JS;
 XX
 DR WPI; 2001-308082/32.
 DR N-PSDB; AAP85636.
 XX
 PT New vitamin D response element-binding protein (VDRE-BP) useful in
 PT modifying vitamin D receptor activity, in producing anti-VDRE-BP
 PT antibodies, in identifying agonists and antagonists of the protein, or in
 PT gene therapy.
 XX
 PS Claim 17; Page 71-72; 81pp; English.
 XX
 CC The present invention provides the protein and coding sequences for two
 CC vitamin D response element binding proteins (VDRE-BPs) from the marmoset.
 CC These confer vitamin D resistance on the primate, a characteristic which
 CC is associated with high circulating levels of other steroid hormones. The
 CC sequences provided by the invention can be used to identify treatments
 CC for osteoporosis, hypercalcaemia, vitamin D intoxication, steroid hormone
 CC hypersecretion, glucocorticoid mediated disorders, lymphoma and granuloma
 CC forming diseases. The present sequence is one of the VDRE-BPs of the
 CC invention
 XX
 SQ Sequence 320 AA;
 Query Match 93.8%; Score 228; DB 4; Length 320;
 Best Local Similarity 100.0%; Pred. No. 2.7e-18;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NOSSNFGPMKGNFGRRSSGPGGGGQYFAKPRNQGGYGG 40
 DB 268 NOSSNFGPMKGNFGRRSSGPGGGGQYFAKPRNQGGYGG 307
 RESULT 14
 ABB00955
 ID ABB00955 standard; protein; 320 AA.
 XX
 AC ABB00955;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #946.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US008631.
 XX

```

PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
PA (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX DR N-PSDB; AAS65142.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX Claim 20; SEQ ID NO 31314; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (I) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful in medical imaging
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX involving aberrant protein expression or biological activity. The
XX polypeptide and polynucleotide sequences have application in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
XX amino acid sequences of the invention. Note: The sequence data for this
XX patent did not appear in the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX Sequence 320 AA;
SQ
Query Match 93.8%; Score 228; DB 4; Length 320;
Best Local Similarity 100.0%; Pred. No. 2.7e-18;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NQSNFGPMKGNFGGRSGPYGGGQYFAKPRNQGYYG 40
DB 268 NQSNFGPMKGNFGGRSGPYGGGQYFAKPRNQGYYG 307
RESULT 15
ABBS7241
ID ABB57241 standard; protein; 320 AA.
XX AC ABB57241;
XX DT 07-MAR-2002 (first entry)
XX DE Mouse ischaemic condition related protein sequence SEQ ID NO:645.
XX KW Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
XX KV vasospastic ischaemia; ischaemic condition; ischaemic disease.
XX OS Mus musculus.
XX PN WO200118188-A2.
XX PD 22-NOV-2001.
XX PF 18-MAY-2001; 2001WO-JP004192.
XX PR 18-MAY-2000; 2000JP-00145977.
XX

```

```

PA (UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.
XX Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
XX WPI; 2002-034733/04.
XX DR N-PSDB; ABI99624.
XX PT Examining the ischemic condition (e.g. occlusive ischemia) by measuring
PT expression levels of particular genes defined in the specification or by
PT determining the expression profile of a gene group comprising these
PT genes.
XX Claim 2; Page 1623-1625; 2690pp; English.
XX The present invention describes a method for examining ischaemic
XX conditions, comprising measuring the expression levels of particular
XX genes (I) in a test sample or determining the expression profile of a
XX gene group in the sample comprising genes selected from (I). The method
XX is useful for examining the ischaemic condition (e.g. compressive
XX ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
XX expression levels of particular genes (ABI99202 to ABI99912, encoding the
XX protein sequences in ABB57020 to ABB57374) or by determining the
XX expression profile of a gene group comprising these genes. The expression
XX levels or expression profiles produced by these genes are used as an
XX indicator when screening for ischaemic condition-improving drugs or
XX therapeutics for ischaemic diseases. ABI99913 and ABI99914 represent PCR
XX primers for a mouse ischaemic condition related sequence, which are used
XX in the exemplification of the present invention
XX Sequence 320 AA;
SQ
Query Match 93.8%; Score 228; DB 5; Length 320;
Best Local Similarity 100.0%; Pred. No. 2.7e-18;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NQSNFGPMKGNFGGRSGPYGGGQYFAKPRNQGYYG 40
DB 268 NQSNFGPMKGNFGGRSGPYGGGQYFAKPRNQGYYG 307
Search completed: October 22, 2004, 13:43:04
Job time : 161 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 22, 2004, 13:40:25 ; Search time 37 seconds

(without alignments)

109.219 Million cell updates/sec

Title: US-09-763-982B-1

Perfect score: 243

Sequence: 1 NQSNFQPMKGNFGGRSSG.....GGGGQYFAKPRNQGGYGGC 42

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79.*

1: Piri.*

2: Pir2.*

3: Pir3.*

4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	228	93.8	320	1 DDRT	helix-destabilizin
2	228	93.8	320	1 A4485	heterogeneous ribo
3	228	93.8	320	2 S04617	heterogeneous ribo
4	228	93.8	320	2 S02061	heterogeneous ribo
5	222	91.4	320	2 S30192	heterogeneous ribo
6	127	52.3	365	2 A34840	heterogeneous ribo
7	123	50.6	346	2 A30775	ribonucleoprotein
8	118	48.6	373	2 S40777	heterogeneous ribo
9	114.5	47.1	385	2 S40778	ribonucleoprotein
10	114	46.9	353	2 B34504	heterogeneous nucl
11	108	44.4	358	2 S40776	ribonucleoprotein
12	107.5	44.2	351	2 B34840	heterogeneous ribo
13	97	39.9	169	1 S38331	glycine-rich RNA-b
14	93.5	38.5	168	1 S12312	glycine-rich RNA-b
15	90	37.0	165	2 T03583	glycine-rich RNA-b
16	89.5	36.8	156	2 S41771	glycine-rich RNA-b
17	87.5	36.0	365	2 A26459	helix-destabilizin
18	87	35.8	173	2 S53050	RNA binding protei
19	85.5	35.2	157	2 S04536	embryonic abundan
20	85	35.0	127	2 S48195	GCR 20 protein - f
21	84.5	34.8	112	2 T22078	hypothetical prote
22	84.5	34.8	220	2 A44805	eggshell protein p
23	84.5	34.8	259	2 T15126	hypothetical prote
24	84.5	34.8	561	2 A31994	keratin 10, type I
25	84	34.6	688	2 T48796	probable ATP-depen
26	83.5	34.4	269	2 I52962	FBRNP - human
27	83.5	34.4	345	1 B41732	heterogeneous nucl
28	83	34.2	166	2 T10463	glycine-rich prote
29	83	34.2	341	2 T27929	hypothetical prote

30	82.5	34.0	169	2 T10465	glycine-rich prote
31	82.5	34.0	334	2 S53490	RNA-binding protei
32	82.5	34.0	604	2 S13653	ATP-dependent RNA
33	82.5	34.0	633	2 H84854	probable ATP-depen
34	82	33.7	161	2 S74453	glycine-rich RNA-b
35	82	33.7	175	2 S54255	probable glycine r
36	81.5	33.5	290	2 T48274	hypothetical prote
37	81.5	33.5	341	2 B75374	conserved hypothet
38	81.5	33.5	593	1 KRHU0	keratin 10, type I
39	81	33.3	145	2 T01356	glycine-rich RNA b
40	81	33.3	386	1 S22315	snRNP-associated p
41	81	33.3	1585	2 T31611	hypothetical prote
42	80.5	33.1	82	2 S19774	glycine-rich prote
43	80.5	33.1	142	2 S12311	glycine-rich RNA-b
44	80.5	33.1	157	1 S14857	glycine-rich prote
45	80	32.9	155	2 S20846	glycine-rich prote

ALIGNMENTS

RESULT 1

DDRT

helix-destabilizing protein - rat

N:Alternate names: single-stranded DNA-binding protein

C:Species: Rattus norvegicus (Norway rat)

C>Date: 04-Dec-1986 #sequence_revision 04-Dec-1986 #text_change 13-Jun-1997

C:Accession: A02682

R:Cobianchi, F.; SenGupta, D.N.; Zmudzka, B.Z.; Wilson, S.H.

J. Biol. Chem. 261, 3536-3543, 1986

A:Title: Structure of rodent helix-destabilizing protein revealed by cDNA cloning.

A:Reference number: A02682; MUID:86140140; PMID:3005291

A:Accession: A02682

A:Molecule type: mRNA

A:Residues: 1-320 <COB>

C:Comment: This protein was isolated from the brain.

C:Superfamily: helix-destabilizing protein; ribonucleoprotein repeat homology

C:Keywords: acetylated amino end; brain; DNA binding; duplication; methylated amino acid

F:13-81/Domain: ribonucleoprotein repeat homology <RM1>

F:106-172/Domain: ribonucleoprotein repeat homology <RM2>

F:2/Modified site: acetylated amino end (Ser) (in mature form) #status predicted

F:194/Modified site: omega-N,omega-N-dimethylarginine (Arg) #status predicted

Query Match 93.8% Score 228; DB 1; Length 320;

Best Local Similarity 100.0%; Pred. No. 1.6e-17;

Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NQSNFQPMKGNFGGRSSGPGYGGQYFAKPRNQGGYGG 40

Db 268 NQSNFQPMKGNFGGRSSGPGYGGQYFAKPRNQGGYGG 307

RESULT 2

A4485

heterogeneous ribonuclear particle protein A1 - mouse

N:Alternate names: helix-destabilizing protein; hnRNP core protein A1; single stranded D

C:Species: Mus musculus (house mouse)

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004

C:Accession: A44485; JC5070; JC5071

R:Ben-David, Y.; Bani, M.R.; Chabot, B.; De Koven, A.; Bernstein, A.

Mol. Cell. Biol. 12, 4449-4455, 1992

A:Title: Retroviral insertions downstream of the heterogeneous nuclear ribonucleoprotein

A:Reference number: A44485; MUID:93024387; PMID:1406633

A:Accession: A44485

A:Molecule type: mRNA

A:Residues: 1-320 <BEN>

A:Cross-references: UNIPROT:P49312; GB:M99167; NID:G193323; PIDN:AAA37633.1; PID:G193324

A:Experimental source: spleen

A:Note: sequence extracted from NCBI backbone (NCBIP:114165)

R:Onishi, Y.; Kizaki, H.

Biochem. Biophys. Res. Commun. 228, 7-13, 1996

A:Title: Molecular cloning of the genes suppressed in RVC lymphoma cells by topoisomerase

A:Reference number: JC5070; MUID:97069646; PMID:8912629

A:Accession: JC5070
 A:Molecule type: mRNA
 A:Residues: 1-320 <ONI>
 A:Cross-references: DDBJ:D86729; NID:g1711241; PIDN:BAAL1362.1; PID:g1711242
 A:Experimental source: lymphoma cell
 A:Accession: JC5071
 A:Molecule type: mRNA
 A:Residues: 1-251,279-320 <ON2>
 A:Cross-references: DDBJ:D86728
 A:Experimental source: lymphoma cell
 C:Comment: This protein inhibits the activity of the SF2/SAP factor in regulating the splicing of pre-mRNA.
 C:Superfamily: helix-destabilizing protein; ribonucleoprotein repeat homology <RRM1>
 C:Keywords: acetylated amino end; alternative splicing; DNA binding; duplication; methylated
 F:15-81/Domain: ribonucleoprotein repeat homology <RRM1>
 F:106-172/Domain: ribonucleoprotein repeat homology <RRM2>
 F:2/Modified site: acetylated amino end (Ser) (in mature form) #status predicted
 F:2/Modified site: acetylated amino end (Ser) (in mature form) #status predicted
 F:194/Modified site: omega-N,omega-N-dimethylarginine (Arg) #status predicted

Query Match 93.8%; Score 228; DB 1; Length 320;
 Best Local Similarity 100.0%; Pred. No. 1.6e-17;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NQSNFGPMKGNFGGRSSGPGYGGGQYFAKPRNQGGYGG 40
 Db 268 NQSNFGPMKGNFGGRSSGPGYGGGQYFAKPRNQGGYGG 307

RESULT 3
 S04617
 heterogenous ribonuclear particle protein A1 - human
 N:Alternate names: hnRNP core protein A1
 N:Contains: helix-destabilizing protein UP1
 C:Species: Homo sapiens (man)
 C:Date: 31-Mar-1990 #sequence revision 31-Mar-1990 #text_change 23-Jul-1999
 C:Accession: S04617; S00749; A24894; S45143
 R:Biamonti, M.; Buvoili, M.; Bassi, M.T.; Morandi, C.; Cobianchi, F.; Riva, S.
 J. Mol. Biol. 207, 491-503, 1989
 A:Title: Isolation of an active gene encoding human hnRNP protein A1. Evidence for alternative splicing.
 A:Reference number: S04617; MUID:89342435; PMID:2760922
 A:Accession: S04617
 A:Molecule type: DNA
 A:Residues: 1-320 <B1A>
 A:Cross-references: EMBL:X12671; NID:g32344; PIDN:CAA31191.1; PID:g296650
 R:Biamonti, M.; Buvoili, M.; Bassi, M.T.; Ghettti, A.; Riva, S.; Morandi, C.
 Nucleic Acids Res. 16, 3751-3770, 1988
 A:Title: CDNA cloning of human hnRNP protein A1 reveals the existence of multiple mRNA isoforms.
 A:Reference number: S00749; MUID:88233978; PMID:2836799
 A:Accession: S00749
 A:Molecule type: mRNA
 A:Residues: 1-320 <B1A>
 A:Cross-references: EMBL:X06747
 R:Biamonti, M.; Morandi, C.; Tsoulfas, P.; Pandolfo, M.; Biamonti, G.; Merrill, B.; Williams, E.M.B. J. 5, 2267-2273, 1986
 A:Title: Mammalian single-stranded DNA binding protein UP I is derived from the hnRNP core protein A1.
 A:Reference number: A24894; MUID:87053868; PMID:3023065
 A:Accession: A24894
 A:Molecule type: mRNA
 A:Residues: 125-139, P, 141-320 <RIV>
 A:Cross-references: GB:X04347; NID:g37609; PIDN:CAA27874.1; PID:g37610
 A:Note: The authors translated the codon CCA for residue 140 as Arg.
 R:Knudsen, S.M.; Jeffers, H.
 Submitted to the EMBL Data Library, June 1994
 A:Description: Cloning and sequencing of a splice variant of human hnRNP core protein A1.
 A:Reference number: S45143
 A:Accession: S45143
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-320 <RNU>
 A:Cross-references: EMBL:X79536; NID:g456897; PIDN:CAA56072.1; PID:g456898
 C:Genetics:
 A:Gene: GDB:HNRP1
 A:Cross-references: GDB:127389; OMIM:164017
 A:Map position: 12q13-12q13

A:Introns: 5/3; 44/3; 93/3; 164/1; 195/1; 226/1; 251/1; 303/1
 C:Superfamily: helix-destabilizing protein; ribonucleoprotein repeat homology
 C:Keywords: acetylated amino end; alternative splicing; DNA binding; duplication; methylated
 F:15-81/Domain: ribonucleoprotein repeat homology <RRM1>
 F:106-172/Domain: ribonucleoprotein repeat homology <RRM2>
 F:2/Modified site: acetylated amino end (Ser) (in mature form) #status predicted
 F:2/Modified site: omega-N,omega-N-dimethylarginine (Arg) #status experimental
 F:194/Modified site: omega-N,omega-N-dimethylarginine (Arg) #status predicted
 F:206,218,225,232/Modified site: omega-N,omega-N-dimethylarginine (Arg) #status predicted

Query Match 93.8%; Score 228; DB 2; Length 320;
 Best Local Similarity 100.0%; Pred. No. 1.6e-17;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NQSNFGPMKGNFGGRSSGPGYGGGQYFAKPRNQGGYGG 40
 Db 268 NQSNFGPMKGNFGGRSSGPGYGGGQYFAKPRNQGGYGG 307

RESULT 4
 S02061
 heterogenous ribonuclear particle protein A1.beta - human
 N:Alternate names: helix-destabilizing protein; hnRNP core protein A1; single stranded DN
 C:Species: Homo sapiens (man)
 C:Date: 31-Dec-1990 #sequence revision 31-Dec-1990 #text_change 09-Jul-2004
 C:Accession: S02061
 R:Biamonti, M.; Biamonti, G.; Tsoulfas, P.; Bassi, M.T.; Ghettti, A.; Riva, S.; Morandi, C.
 Nucleic Acids Res. 16, 3751-3770, 1988
 A:Title: CDNA cloning of human hnRNP protein A1 reveals the existence of multiple mRNA isoforms.
 A:Reference number: S00749; MUID:88233978; PMID:2836799
 A:Accession: S02061
 A:Molecule type: mRNA
 A:Residues: 1-320 <BUV>
 A:Cross-references: UNIPROT:P09651; EMBL:X06747; NID:g36101; PIDN:CAA29922.1; PID:g36102
 C:Superfamily: helix-destabilizing protein; ribonucleoprotein repeat homology
 C:Keywords: acetylated amino end; alternative splicing; DNA binding; methylated amino acid
 F:15-81/Domain: ribonucleoprotein repeat homology <RRM1>
 F:106-172/Domain: ribonucleoprotein repeat homology <RRM2>
 F:2/Modified site: acetylated amino end (Ser) (in mature form) #status predicted
 F:194/Modified site: omega-N,omega-N-dimethylarginine (Arg) #status predicted

Query Match 93.8%; Score 228; DB 2; Length 320;
 Best Local Similarity 100.0%; Pred. No. 1.6e-17;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NQSNFGPMKGNFGGRSSGPGYGGGQYFAKPRNQGGYGG 40
 Db 268 NQSNFGPMKGNFGGRSSGPGYGGGQYFAKPRNQGGYGG 307

RESULT 5
 S30192
 heterogenous ribonuclear particle protein A1 - rhesus macaque
 N:Alternate names: heterogenous nuclear ribonucleoprotein A1; hnRNP A1
 C:Species: Macaca mulatta (rhesus macaque)
 C:Date: 03-May-1994 #sequence revision 20-Feb-1995 #text_change 09-Jul-2004
 C:Accession: S30192; S34126
 R:An, G.; Wu, R.
 Biochim. Biophys. Acta 1172, 292-300, 1993
 A:Title: CDNA cloning of a hnRNP A1 isoform and its regulation by retinol in monkey tract
 A:Reference number: S30192; MUID:93192320; PMID:8448206
 A:Accession: S30192
 A:Molecule type: mRNA
 A:Residues: 1-320 <ANG>
 A:Cross-references: UNIPROT:Q28521; EMBL:M84334; NID:g1339834; PIDN:AA801436.1; PID:g1339834
 R:An, G.; Wu, R.
 submitted to the EMBL Data Library, February 1992
 A:Reference number: S34126
 A:Accession: S34126
 A:Molecule type: mRNA
 A:Residues: 1C, 3-320 <AN2>
 A:Cross-references: EMBL:M84334
 C:Superfamily: helix-destabilizing protein; ribonucleoprotein repeat homology
 C:Keywords: acetylated amino end; methylated amino acid

F:15-81/Domain: ribonucleoprotein repeat homology <RRM1>
 F:106-172/Domain: ribonucleoprotein repeat homology <RRM2>
 F:2/Modified site: acetylated amino end (Ser) (in mature form) #status predicted
 F:194/Modified site: omega-N,omega-N-dimethylarginine (Arg) #status predicted

Query Match 91.4%; Score 222; DB 2; Length 320;
 Best Local Similarity 97.5%; Pred. No. 7e-17; Indels 0; Gaps 0;
 Matches 39; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NOSSNFGPMKGNFGRRSSGPGYGGGQYFAKPRNQGGYGG 40
 |||||
 DB 268 NOSSNFGPMKGNFGRRSSGPGYGGGQYFAKPRNQGGYGG 307
 |||||

RESULT 6
 A34840
 heterogenous ribonuclear particle protein A1.a - African clawed frog
 N:Alternate names: heterogeneous nuclear ribonucleoprotein XAla
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 13-Jul-1990 #sequence_revision 13-Jul-1990 #text_change 09-Jul-2004
 C:Accession: A34840
 R:Kay, B.K.; Sawhney, R.K.; Wilson, S.H.
 Proc. Natl. Acad. Sci. U.S.A. 87, 1367-1371, 1990
 A:Title: Potential for two isoforms of the A1 ribonucleoprotein in Xenopus laevis.
 A:Reference number: A34840; MUID:90160329; PMID:2137612
 A:Accession: A34840
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-365 <XAY>
 A:Cross-references: UNIPROT:P17130; GB:M31041; NID:G214238; PIDN:AAA49741.1; PID:G214239
 C:Superfamily: helix-stabilizing protein; ribonucleoprotein repeat homology
 C:Keywords: alternative splicing; DNA binding; duplication; nucleus
 F:15-81/Domain: ribonucleoprotein repeat homology <RRM1>
 F:106-172/Domain: ribonucleoprotein repeat homology <RRM2>

Query Match 52.3%; Score 127; DB 2; Length 365;
 Best Local Similarity 64.3%; Pred. No. 1.4e-06;
 Matches 27; Conservative 3; Mismatches 6; Indels 3; Gaps 3;

QY 3 SSNFGPMKGNF-GRRSSGPGYGG--GGQYFAKPRNQGGYGGG 41
 |||||
 DB 324 SSNFGPMKGNFGRRSSGPGYGGG---GSASSSGYGGG 362
 |||||

RESULT 7
 S40775
 ribonucleoprotein - African clawed frog
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 19-May-1994 #sequence_revision 26-May-1995 #text_change 09-Jul-2004
 C:Accession: S40775
 R:Good, P.J.; Rebbert, M.L.; Dawid, I.B.
 Nucleic Acids Res. 21, 999-1006, 1993
 A:Title: Three new members of the RNP protein family in Xenopus.
 A:Reference number: S40774; MUID:93197168; PMID:8451200
 A:Accession: S40775
 A:Status: preliminary; nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-346 <GOO>
 A:Cross-references: UNIPROT:P51989; EMBL:L02954; NID:G214740; PIDN:AAA49948.1; PID:G2147
 C:Superfamily: helix-stabilizing protein; ribonucleoprotein repeat homology
 F:10-76/Domain: ribonucleoprotein repeat homology <RRM1>
 F:101-167/Domain: ribonucleoprotein repeat homology <RRM2>

Query Match 50.6%; Score 123; DB 2; Length 346;
 Best Local Similarity 64.4%; Pred. No. 3.7e-06;
 Matches 29; Conservative 1; Mismatches 9; Indels 6; Gaps 3;

QY 2 QSSNFGPMK--GNFGRRSS--GPYGGG--GQYFAKPRNQGGYGG 40
 |||||
 DB 298 QSSNFGPMKGGNFGRRSSGPGYGGGYPGNASGGNGGGYGG 342
 |||||

RESULT 8

S40777
 heterogenous ribonuclear particle protein A3 - African clawed frog
 N:Alternate names: heterogeneous nuclear ribonucleoprotein A3
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 31-Dec-1993 #sequence_revision 02-Aug-1994 #text_change 09-Jul-2004
 C:Accession: S40777; S41738
 R:Good, P.J.; Rebbert, M.L.; Dawid, I.B.
 Nucleic Acids Res. 21, 999-1006, 1993
 A:Title: Three new members of the RNP protein family in Xenopus.
 A:Reference number: S40774; MUID:93197168; PMID:8451200
 A:Accession: S40777
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-373 <GOO1>
 A:Cross-references: UNIPROT:P51968; EMBL:L02956
 R:Good, P.J.; Lai, M.; Rebbert, M.L.; Dawid, I.B.
 submitted to the EMBL Data Library, October 1992
 A:Reference number: S41738
 A:Accession: S41738
 A:Molecule type: mRNA
 A:Residues: 1-85 'S', 87-346, 'S', 348-373 <GOO2>
 A:Cross-references: EMBL:L02956; NID:G214744; PIDN:AAA49949.1; PID:G214745
 C:Superfamily: helix-stabilizing protein; ribonucleoprotein repeat homology
 F:28-94/Domain: ribonucleoprotein repeat homology <RRM1>
 F:119-185/Domain: ribonucleoprotein repeat homology <RRM2>

Query Match 48.6%; Score 118; DB 2; Length 373;
 Best Local Similarity 55.6%; Pred. No. 1.4e-05;
 Matches 25; Conservative 2; Mismatches 6; Indels 12; Gaps 2;

QY 2 QSSNFGPMKGNF-----GRRSSGPGYGGGQYFAKPRNQGGYGGG 41
 |||||
 DB 327 QSSNFGPMKGSFGRSSGRRSGPGYGGG-----YGSGGGGGGG 364
 |||||

RESULT 9
 S40778
 ribonucleoprotein - African clawed frog
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 19-May-1994 #sequence_revision 26-May-1995 #text_change 09-Jul-2004
 C:Accession: S40778
 R:Good, P.J.; Rebbert, M.L.; Dawid, I.B.
 Nucleic Acids Res. 21, 999-1006, 1993
 A:Title: Three new members of the RNP protein family in Xenopus.
 A:Reference number: S40774; MUID:93197168; PMID:8451200
 A:Accession: S40778
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-385 <GOO>
 A:Cross-references: UNIPROT:P51992; EMBL:L02957; NID:G214746; PIDN:AAA49950.1; PID:G21474
 C:Superfamily: helix-stabilizing protein; ribonucleoprotein repeat homology
 F:28-94/Domain: ribonucleoprotein repeat homology <RRM1>
 F:119-185/Domain: ribonucleoprotein repeat homology <RRM2>

Query Match 47.1%; Score 114.5; DB 2; Length 385;
 Best Local Similarity 52.1%; Pred. No. 3.4e-05;
 Matches 25; Conservative 2; Mismatches 10; Indels 11; Gaps 2;

QY 2 QSSNFGPMKGNF-----GRRSSGPGYGGGQYFAKPRNQGGYGGG 41
 |||||
 DB 334 QSSNFGPMKAGNFGRRSTGGSGSGSGPGYGGG---YGSGGGGGGGGG 378
 |||||

RESULT 10
 B34504
 heterogenous nuclear ribonucleoprotein B1 - human
 N:Alternate names: heterogeneous nuclear ribonucleoprotein B1; hnRNP protein B1; NEPK
 C:Contains: heterogeneous nuclear ribonucleoprotein A2
 C:Species: Homo sapiens (man)
 C:Date: 22-Jun-1990 #sequence_revision 22-Jun-1990 #text_change 09-Jul-2004
 C:Accession: A56845; B34504; A34504; S48057; PC2222; B61013; B56845
 R:Kozu, T.; Henrich, B.; Schaefer, K.P.

R.; Good, P.J.; Rebbert, M.L.; Dawid, I.B.
Nucleic Acids Res. 21, 999-1006, 1993
A>Title: Three new members of the RNP protein family in Xenopus.
A:Reference number: S40774; MUID:93197168; PMID:8451200
A:Accession: S40776
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: mRNA
A:Residues: 1-358 <GO>
A:Cross-references: UNIPROT:P51990; EMBL:L02955; NID:g214742; PIDN:AAB59951.1; PID:g214742
A>Note: The nucleotide sequence was submitted to the EMBL Data Library, October 1992
C:Superfamily: helix-destabilizing protein; ribonucleoprotein repeat homology
F:107-76/Domain: ribonucleoprotein repeat homology <RRM1>
F:101-167/Domain: ribonucleoprotein repeat homology <RRM2>

Query Match 44.4%; Score 108; DB 2; Length 358;
Best Local Similarity 56.2%; Pred. No. 0.00016;
Matches 27; Conservative 2; Mismatches 7; Indels 12; Gaps 4;

QY 2 QSSNGPMPK-GGNFGGRSS--GPYGGGGQYFAKPRN-----QGCGYG 40
DB 310 QSSSYGPMPKSGGNFGGRNSMGSPGGGNY---GPGNSGASGCGGYGG 354
|||||

RESULT 12
B34840
heterogeneous ribonuclear particle protein Al.b - African clawed frog
N:Alternate names: heterogeneous nuclear ribonucleoprotein XAlb
C:Species: Xenopus laevis (African clawed frog)
C>Date: 13-Jul-1990 #sequence_revision 13-Jul-1990 #text_change 09-Jul-2004
C:Accession: B34840
R:Kay, B.K.; Sawhney, R.K.; Wilson, S.H.
Proc. Natl. Acad. Sci. U.S.A. 87, 1367-1371, 1990
A>Title: Potential for two isoforms of the Al ribonucleoprotein in Xenopus laevis.
A:Reference number: A34840; MUID:90160329; PMID:2137612
A:Accession: B34840
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-351 <KAY>
A:Cross-references: UNIPROT:P17130; GB:M30575; NID:g214240; PIDN:AAA49742.1; PID:g214241
C:Superfamily: helix-destabilizing protein; ribonucleoprotein repeat homology
C:Keywords: alternative splicing; DNA binding; duplication; nucleus
F:15-81/Domain: ribonucleoprotein repeat homology <RRM1>
F:106-172/Domain: ribonucleoprotein repeat homology <RRM2>

Query Match 44.2%; Score 107.5; DB 2; Length 351;
Best Local Similarity 87.0%; Pred. No. 0.00017;
Matches 20; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

QY 3 SSNGPMPKGNF-GGRSGPYGG 24
DB 324 SSNGPMPKGNYGGRNSGPYGG 346
|||||

RESULT 13
S38331
glycine-rich RNA-binding protein - rape
C:Species: Brassica napus (rape)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: S38331; S25120
R:Bergeton, D.; Beaueigle, D.; Bellemare, G.
Biochim. Biophys. Acta 1216, 123-125, 1993
A>Title: Sequence and expression of a gene encoding a protein with RNA-binding and glycer
A:Reference number: S38331; MUID:94032471; PMID:7916642
A:Accession: S38331
A:Molecule type: DNA
A:Residues: 1-169 <BER>
A:Cross-references: UNIPROT:Q05966; EMBL:Z14143; NID:g17818; PIDN:CAA78513.1; PID:g17819
C:Genetics:
A:Introns: 36/3
C:Superfamily: glycine-rich RNA-binding protein; ribonucleoprotein repeat homology
F:17-74/Domain: ribonucleoprotein repeat homology <RRM1>

Query Match 39.9%; Score 97; DB 1; Length 169;

Best Local Similarity 51.4%; Pred. No. 0.0012;
Matches 18; Conservative 2; Mismatches 15; Indels 0; Gaps 0;

QY 7 GPMKGNFGSRSSGPGYGGGQYFAKPRNQGGYGGG 41
DB 91 GORGGGYGRGGGGYGGGGYGRGGGGYGGG 125

RESULT 14
S12312
glycine-rich RNA-binding protein (clone S2) - sorghum
C:Species: Sorghum bicolor (sorghum)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: S12312
R:Cretin, C.; Puigdomenech, P.
Plant Mol. Biol. 15, 783-785, 1990
A:Title: Glycine-rich RNA-binding proteins from Sorghum vulgare.
A:Reference number: S12311; MUID:91346715; PMID:1715211
A:Accession: S12312
A:Molecule type: mRNA
A:Residues: 1-168 <CR>
A:Cross-references: UNIPROT:Q99070; EMBL:X57662; NID:g21624; PIDN:CAA40862.1; PID:g21625
A>Note: In the authors' translation two additional Gly are shown after 110-Gly
C:Superfamily: glycine-rich RNA-binding protein; ribonucleoprotein repeat homology
C:Keywords: GTP binding
F:9-76/Domain: ribonucleoprotein repeat homology <RRM1>

Query Match 38.5%; Score 93.5; DB 1; Length 168;
Best Local Similarity 56.2%; Pred. No. 0.0028;
Matches 18; Conservative 2; Mismatches 11; Indels 1; Gaps 1;

QY 11 GGNFGRSSGPY-GGGGQYFAKPRNQGGYGGG 41
DB 100 GGGYGRGGYGGGGYGGGGYGRGGGGYGGG 131

RESULT 15
T03583
glycine-rich RNA-binding protein - rice
C:Species: Oryza sativa (rice)
C>Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 09-Jul-2004
C:Accession: T03583
R:Lee, M.C.; Kim, C.S.; Yi, B.Y.; Lee, J.S.; Eun, M.Y.
submitted to the EMBL Data Library, June 1997
A:Description: Isolation and characterization of RNA-binding glycine rich protein of rice
A:Reference number: Z14958
A:Accession: T03583
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-165 <LEE>
A:Cross-references: UNIPROT:O24184; EMBL:AF009411; NID:g2267592; PIDN:AAB63589.1; PID:g2
A:Experimental source: cv. Milyang 23
C:Superfamily: glycine-rich RNA-binding protein; ribonucleoprotein repeat homology
F:9-76/Domain: ribonucleoprotein repeat homology <RRM>

Query Match 37.0%; Score 90; DB 2; Length 165;
Best Local Similarity 58.1%; Pred. No. 0.0066;
Matches 18; Conservative 2; Mismatches 9; Indels 2; Gaps 1;

QY 11 GGNFGRSSGPYGGGQYFAKPRNQGGYGGG 41
DB 103 GGGYGRGGYGGGGYGGGGY--GQREGGYGGG 131

Search completed: October 22, 2004, 13:47:49
Job time : 38 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 22, 2004, 13:37:19 ; Search time 194 Seconds
(without alignments)
124.566 Million cell updates/sec

Title: US-09-763-982B-1

Perfect score: 243

Sequence: 1 NQSNFGPMKGNFGGRSSG.....GGGGQYFAKPRNQGGYGGC 42

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 02:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	228	93.8	298	Q9BSM5	Q9bsm5 homo sapien
2	228	93.8	313	Q6P6G9	Q6p6g9 rattus norv
3	228	93.8	313	AAH62235	Aah62235 rattus norv
4	228	93.8	319	1 ROAL_MOUSE	P49312 mus musculus
5	228	93.8	319	1 ROAL_RAT	P04256 rattus norv
6	228	93.8	320	Q6IPF2	Q6ipf2 homo sapien
7	228	93.8	320	AAH12158	Aah12158 homo sapi
8	228	93.8	320	AAH33714	Aah33714 homo sapi
9	228	93.8	320	AAH02355	Aah02355 homo sapi
10	228	93.8	320	AAH09600	Aah09600 homo sapi
11	228	93.8	320	AAH70315	Aah70315 homo sapi
12	228	93.8	320	AAH71945	Aah71945 homo sapi
13	228	93.8	371	1 ROAL_HUMAN	P09651 homo sapien
14	222	91.4	319	1 ROAL_MACMU	Q28521 macaca mula
15	219	90.1	70	2 Q9H4S1	Q9h4s1 homo sapien
16	127	52.3	365	1 ROAL_XENLA	P17130 xenopus lae
17	127	52.3	365	2 Q6I2Z9	Q6i2z9 xenopus lae
18	127	52.3	365	2 AAH2090	Aah2090 xenopus lae
19	125	51.4	63	2 Q8NFG3	Q8nfg3 homo sapien
20	125	51.4	193	2 Q70592	Q70592 rattus norv
21	125	51.4	318	2 Q6P6I7	Q6p6i7 mus musculus
22	125	51.4	318	2 AAH62198	Aah62198 mus muscu
23	125	51.4	357	2 AAQ63631	Aaq63631 rattus norv
24	125	51.4	378	1 ROA3_HUMAN	P51991 homo sapien
25	125	51.4	378	2 AAQ63629	Aaq63629 homo sapi
26	125	51.4	379	1 ROA3_MOUSE	Q8b905 mus musculus
27	125	51.4	379	1 ROA3_RAT	Q6urk4 rattus norv
28	125	51.4	379	2 AAQ63630	Aaq63630 rattus norv
29	125	51.4	379	2 AAH64824	Aah64824 mus muscu
30	123	50.6	346	1 RO21_XENLA	P51989 xenopus lae
31	120	49.4	388	2 Q7SXQ3	Q7sxq3 brachydanio

32 117 48.1 314 2 Q6NYW8
33 117 48.1 314 2 AAH66434
34 117 48.1 346 2 Q7ZWV4
35 116 47.7 373 1 RO31_XENLA
36 114.5 47.1 385 1 RO32_XENLA
37 114 46.9 219 2 Q6PCV9
38 114 46.9 219 2 AAH59107
39 114 46.9 301 2 Q8C2A0
40 114 46.9 341 1 ROA2_MOUSE
41 114 46.9 341 2 Q9TTV2
42 114 46.9 341 2 Q912R9
43 114 46.9 353 1 ROA2_HUMAN
44 114 46.9 353 2 Q8CJ71
45 112.5 46.3 422 2 Q803K3

Q6nyw8 brachydanio
Aah66434 brachydan
Q7zwv4 xenopus lae
P51968 xenopus lae
P51932 xenopus lae
Q6pcv9 mus musculus
Aah59107 mus muscu
Q8c2a0 mus musculu
Q88569 mus musculu
Q9ttv2 saguinus oe
Q912r9 mus musculu
P22626 homo sapien
Q8cj71 mus musculu
Q803k3 brachydanio

ALIGNMENTS

RESULT 1
Q9BSM5 PRELIMINARY; PRT; 298 AA.
ID Q9BSM5;
AC Q9BSM5;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=skin;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC004945; AAH04945.1; -.
DR InterAct; Q9BSM5; -.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00076; RRM_1; 2.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS50102; RRM; 2.
DR PROSITE; PS00030; RRM_RNP_1; 2.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 298 AA; 31752 MW; 8A4CDFCB838D78BB CRC64;
Query Match 93.8%; Score 228; DB 2; Length 298;
Best Local Similarity 100.0%; Pred. No. 2.1e-16;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NQSNFGPMKGNFGGRSSGPGYGGGQYFAKPRNQGGYGG 40
DB 246 NQSNFGPMKGNFGGRSSGPGYGGGQYFAKPRNQGGYGG 285
RESULT 2
Q6P6G9 PRELIMINARY; PRT; 313 AA.
ID Q6P6G9;
AC Q6P6G9;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Hnrp1 protein.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pituitary gland;
RC MEDLINE=22388257; PubMed=12477932;
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.D., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshitsuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Pahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pituitary gland;
 RA Strausberg R.;
 DR EMBL: BC062235; AAH62235.1; - - - - -
 DR InterPro: IPR000504; RNA_rec_mot.
 DR Pfam: PF00076; RRM_1; 2.
 DR SMART: SMO0360; RRM; 2.
 DR PROSITE: PS0102; RRM; 2.
 DR PROSITE: PS0030; RRM_RNP_1; UNKNOWN 2.
 SQ SEQUENCE 313 AA; 33620 MW; AE0EEC10972DEF7D CRC64;

Query Match 93.8%; Score 228; DB 2; Length 313;
 Best Local Similarity 100.0%; Pred. No. 2.2e-16;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NQSNFPGMKGNFGGRSSGPGYGGGQYFAKPRNQGYYG 40
 |||||
 DB 261 NQSNFPGMKGNFGGRSSGPGYGGGQYFAKPRNQGYYG 300
 |||||

RESULT 3
 AAH62235 PRELIMINARY; PRT; 313 AA.
 ID AAH62235
 AC AAH62235
 DT 02-MAR-2004 (TrEMBLrel. 27, Created)
 DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
 DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hirpal protein.
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 CX NCBI_TaxID=10116;
 [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pituitary gland;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.D., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshitsuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Pahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,

RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pituitary gland;
 RA Strausberg R.;
 DR EMBL: BC062235; AAH62235.1; - - - - -
 DR EMBL: BC062235; AAH62235.1; - - - - -
 SQ SEQUENCE 313 AA; 33620 MW; AE0EEC10972DEF7D CRC64;
 Query Match 93.8%; Score 228; DB 2; Length 313;
 Best Local Similarity 100.0%; Pred. No. 2.2e-16;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NQSNFPGMKGNFGGRSSGPGYGGGQYFAKPRNQGYYG 40
 |||||
 DB 261 NQSNFPGMKGNFGGRSSGPGYGGGQYFAKPRNQGYYG 300
 |||||

RESULT 4

ROAL MOUSE
 ID ROAL MOUSE STANDARD; PRT; 319 AA.
 AC P49312; P97312;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Heterogeneous nuclear ribonucleoprotein A1 (Helix-destabilizing
 DE protein) (Single-strand binding protein) (hnRNP core protein A1) (HDP-
 DE 1) (Topoisomerase-inhibitor suppressed).
 GN Name=HnRNP; Synonyms=Fli-2, Tis;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CX NCBI_TaxID=10090;
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93024387; PubMed=1406633;
 RA Ben-David Y., Bani M.R., Chabot B., de Koven A., Bernstein A.;
 RT "Retroviral insertions downstream of the heterogeneous nuclear
 RT ribonucleoprotein A1 gene in erythroleukemia cells: evidence that A1
 RT is not essential for cell growth."
 RL Mol. Cell. Biol. 12:4449-4455 (1992).
 [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lymphoma;
 RX MEDLINE=97069646; PubMed=8912629;
 RA Onishi Y., Kizaki H.;
 RT "Molecular cloning of the genes suppressed in RVC lymphoma cells by
 RT topoisomerase inhibitors."
 RL Biochem. Biophys. Res. Commun. 228:7-13 (1996).
 [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Pancreas, and Thymus;
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nikaide I., Osato N., Saito R., Suzuki H., Yamana I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
 RA Baldairelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.B., Cousins S.,
 RA Dalla B., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Gasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Naglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sandelin A., Schneider C., Sempke C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tonita M.,

Verardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C., Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L., Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N., Hirozaki-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K., Shiroaki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S., Hara A., Hashizume W., Inotani K., Ishii Y., Itoh M., Kagawa I., Miyazaki A., Sakai K., Sasaki D., Shibata K., Shingawa A., Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J., Birney E., Hayashizaki Y.;
 "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs";
 Nature 420:563-573 (2002).
 [4]
 RN PHOSPHORYLATION SITES SER-5 AND SER-312.
 RP Pubmed=14729942; DOI=10.1074/mcp.D300003-MCP200;
 RX Shu H., Chen S., Bi Q., Mumbay M., Breken D.L.;
 RT "Identification of phosphoproteins and their phosphorylation sites in the WEHI-231 B lymphoma cell line.";
 RL Mol. Cell. Proteomics 3:279-286(2004).
 CC -!- FUNCTION: Involved in the packaging of pre-mRNA into hnRNP particles, transport of poly(A) mRNA from the nucleus to the cytoplasm and may modulate splice site selection.
 CC -!- SUBCELLULAR LOCATION: Nuclear. Shuttles continuously between the nucleus and the cytoplasm along with mRNA. Component of ribonucleosomes.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=Long;
 CC IsoId=P49312-1; Sequence=Displayed;
 CC Name=Short;
 CC IsoId=P49312-2; Sequence=VSP_005825;
 CC -!- SIMILARITY: Contains 2 RNA recognition motif (RRM) domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).
 CC EMBL; M99167; AAA37633.1; -;
 CC EMBL; D86729; BAA13162.1; -;
 CC EMBL; D86728; BAA13161.1; -;
 CC EMBL; AK007802; BAB25267.1; -;
 CC EMBL; AK089308; BAC40273.1; -;
 CC FIR; A44485; A44485.
 CC HSSP; P09651; 2UPI.
 CC MGD; MGI:104820; Hnrp1.
 CC InterPro; IPR000504; RNA_rec_mot.
 CC Pfam; PF000076; RRM_1; 2.
 CC PROSITE; PS50102; RRM; 2.
 CC PROSITE; PS00030; RRM_RNP_1; 2.
 CC Alternative splicing; Methylation; mRNA transport; Nuclear protein;
 CC Phosphorylation; Repeat; Ribonucleoprotein; RNA-binding; Transport.
 CC INIT MET 0 0
 CC By similarity.
 CC DOMAIN 3 93
 CC Globular A domain.
 CC DOMAIN 94 184
 CC Globular B domain.
 CC DOMAIN 13 96
 CC RNA-binding (RRM) 1.
 CC DOMAIN 104 183
 CC RNA-binding (RRM) 2.
 CC DOMAIN 217 239
 CC RNA-binding RGG-box.
 CC DOMAIN 194 319
 CC Gly-rich.
 CC DOMAIN 267 304
 CC NUCLEAR TARGETING SEQUENCE (BY SIMILARITY).
 CC DOMAIN 307 312
 CC Poly-Ser.
 CC MOD_RES 5 5
 CC Phosphoserine.
 CC FT 193 193
 CC Asymmetric dimethylarginine (By similarity).
 CC MOD_RES 312 312
 CC Phosphoserine.
 CC FT 251 302
 CC Missing (in isoform Short).
 CC VARSPLIC /FTID=VSP_005825.
 CC SEQUENCE 319 AA; 34065 MW; D80974FF2006B303 CRC64;
 CC

```

Query Match          93.8%; Score 228; DB 1; Length 319;
Best Local Similarity 100.0%; Pred. No. 2.2e-16;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  NQSSNFGPMKGGNFGGRSSGPGYGGGQYFAKPRNQGGYGG 40
      |||
DB      267  NQSSNFGPMKGGNFGGRSSGPGYGGGQYFAKPRNQGGYGG 306
      |||

RESULT 5
ROAL_RAT
ID      ID      STANDARD;      PRT;      319 AA.
AC      P04256;
DT      20-MAR-1987 (Rel. 04, Created)
DT      01-MAR-1989 (Rel. 10, Last sequence update)
DT      05-MAR-2004 (Rel. 44, Last annotation update)
DE      Heterogeneous nuclear ribonucleoprotein A1 (Helix-destabilizing
DE      protein) (Single-strand binding protein) (hnRNP core protein A1)
DE      (HDP).
GN      Name=HnRnpA1;
OS      Rattus norvegicus (Rat).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathia; Muridae; Murinae; Rattus.
OC      NCBI_TaxId=10116;
[1]
RN      RNP      SEQUENCE FROM N.A.
RC      TISSUE=Brain;
RX      MEDLINE=88087168; PubMed=3005291;
RA      Cobianchi F., Sengupta D.N., Zmudzka B.Z., Wilson S.H.;
RT      "Structure of rodent helix-destabilizing protein revealed by cDNA
RT      cloning.";
RL      J. Biol. Chem. 261:3536-3543 (1986).
[2]
RN      RNP      REVISION TO 181.
RA      Cobianchi F.;
RL      Submitted (MAY-1986) to the EMBL/GenBank/DBJ databases.
[3]
RN      RNP      PARTIAL SEQUENCE.
RX      MEDLINE=88087168; PubMed=2447078;
RA      Cobianchi F., Karpel R.L., Williams K.R., Notario V., Wilson S.H.;
RT      "Mammalian heterogeneous nuclear ribonucleoprotein complex protein A1.
RT      Large-scale overproduction in Escherichia coli and cooperative binding
RT      to single-stranded nucleic acids.";
RL      J. Biol. Chem. 263:1063-1071 (1988).
CC      -1- FUNCTION: Involved in the packaging of pre-mRNA into hnRNP
CC      particles, transport of poly(A) mRNA from the nucleus to the
CC      cytoplasm and may modulate splice site selection.
CC      -1- SUBCELLULAR LOCATION: Nuclear. Shuttles continuously between the
CC      nucleus and the cytoplasm along with mRNA. Component of
CC      ribonucleosomes.
CC      -1- SIMILARITY: Contains 2 RNA recognition motif (RRM) domains.
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announcement
CC      or send an email to license@isb-sib.ch).
-----
DR      EMBL; M12156; AAA41314.1; -
DR      HSSP; P09651; 1UP1.
DR      RGD; 52334; HnRnpA1.
DR      InterPro; IPR000504; RNA_rec_mot.
DR      Pfam; PF00076; RRM_1; 2.
DR      PROSITE; PS50102; RRM; 2.
DR      PROSITE; PS00030; RRM_RNP_1; 2.
KW      Direct protein sequencing; Methylation; mRNA transport;
KW      Nuclear protein; Phosphorylation; Repeat; Ribonucleoprotein;
KW      RNA-binding; Transport.
FT      INIT MET      0
FT      DOMAIN      3  93      Globular A domain.
FT      DOMAIN      94  184      Globular B domain.

```

FT DOMAIN 13 96 RNA-binding (RRM) 1.
 FT DOMAIN 104 183 RNA-binding (RRM) 2.
 FT DOMAIN 217 239 RNA-binding RGG-box.
 FT DOMAIN 194 319 Gly-rich.
 FT DOMAIN 267 304 NUCLEAR TARGETING SEQUENCE (BY SIMILARITY).
 FT DOMAIN 307 312 Poly-Ser.
 FT MOD_RES 5 5 Phosphoserine (By similarity).
 FT MOD_RES 193 193 Asymmetric dimethylarginine.
 FT MOD_RES 312 312 Phosphoserine (By similarity).
 FT CONFLICT 8 E -> G (in Ref. 3).
 SQ SEQUENCE 319 AA; 34081 MW; CSBEYD183456B303 CRC64;

Query Match 93.8%; Score 228; DB 1; Length 319;
 Best Local Similarity 100.0%; Pred. No. 2.2e-16;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NOSSNFGPMKGNFGGRSSGPGYGGGQYFAKPRNQGYYG 40
 DB 267 NOSSNFGPMKGNFGGRSSGPGYGGGQYFAKPRNQGYYG 306

RESULT 6

OSIPF2 PRELIMINARY; PRT; 320 AA.
 AC Q61PF2;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Heterogeneous nuclear ribonucleoprotein A1, isoform a.
 GN Name=NRPA1; (Human).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 [1]
 RN SEQUENCE FROM N.A.
 RC TISSUE=Skin;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Bosak S.A., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madao A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Skin;
 RA Strausberg R.;
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC071945; AAH71945.1; -;
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF000076; RRM.1; 2.
 DR SMART; SM00360; RRM.2;
 DR PROSITE; PS0102; RRM; 2.
 DR PROSITE; PS00030; RRM_RNP.1; UNKNOWN.2.
 KW Nucleocapsid; Ribonucleoprotein.
 SQ SEQUENCE 320 AA; 34180 MW; 9069C6B408DE1AF3 CRC64;

Query Match 93.8%; Score 228; DB 2; Length 320;
 Best Local Similarity 100.0%; Pred. No. 2.2e-16;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NOSSNFGPMKGNFGGRSSGPGYGGGQYFAKPRNQGYYG 40
 DB 268 NOSSNFGPMKGNFGGRSSGPGYGGGQYFAKPRNQGYYG 307

RESULT 7

AAH12158 PRELIMINARY; PRT; 320 AA.
 ID AAH12158;
 AC AAH12158;
 DT 02-MAR-2004 (TrEMBLrel. 27, Created)
 DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
 DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
 DE Heterogeneous nuclear ribonucleoprotein A1, isoform a.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primata; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 [1]
 RN SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madao A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RA Strausberg R.;
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC012158; AAH12158.1; -;
 KW Nucleocapsid; Ribonucleoprotein.
 SQ SEQUENCE 320 AA; 34196 MW; 59485C9FA1F8A81 CRC64;

Query Match 93.8%; Score 228; DB 2; Length 320;
 Best Local Similarity 100.0%; Pred. No. 2.2e-16;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NOSSNFGPMKGNFGGRSSGPGYGGGQYFAKPRNQGYYG 40
 DB 268 NOSSNFGPMKGNFGGRSSGPGYGGGQYFAKPRNQGYYG 307

RESULT 8

AAH33714 PRELIMINARY; PRT; 320 AA.
 ID AAH33714;
 AC AAH33714;
 DT 02-MAR-2004 (TrEMBLrel. 27, Created)
 DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
 DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
 DE Heterogeneous nuclear ribonucleoprotein A1, isoform a.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

QY 1 NOSSNFGPMKGNFGGRSSGPGYGGGQYFAKPRNQGCGG 40
 DB 268 NOSSNFGPMKGNFGGRSSGPGYGGGQYFAKPRNQGCGG 307

RESULT 11

AAH70315 PRELIMINARY; PRT; 320 AA.
 AC AAH70315;
 DT 13-MAY-2004 (TREMELrel. 27, Created)
 DT 13-MAY-2004 (TREMELrel. 27, Last sequence update)
 DT 13-MAY-2004 (TREMELrel. 27, Last annotation update)
 DE HNRPA1 protein.
 GN HNRPA1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.

RC TISSUE=Bone marrow;
 RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]

RP SEQUENCE FROM N.A.
 RC TISSUE=Bone marrow;
 RX Strausberg R.;

RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC070315; AAH70315.1; -

SQ SEQUENCE 320 AA; 34196 MW; 59485C9F8A1F8AE1 CRC64;

Query Match 93.8%; Score 228; DB 2; Length 320;

Best Local Similarity 100.0%; Pred. No. 2.2e-16; Mismatches 0; Indels 0; Gaps 0;

Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NOSSNFGPMKGNFGGRSSGPGYGGGQYFAKPRNQGCGG 40

DB 268 NOSSNFGPMKGNFGGRSSGPGYGGGQYFAKPRNQGCGG 307

RESULT 12

AAH71945 PRELIMINARY; PRT; 320 AA.
 AC AAH71945;
 DT 01-JUN-2004 (TREMELrel. 27, Created)
 DT 01-JUN-2004 (TREMELrel. 27, Last sequence update)
 DT 01-JUN-2004 (TREMELrel. 27, Last annotation update)
 DE Heterogeneous nuclear ribonucleoprotein A1, isoform A.
 GN HNRPA1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;

RN SEQUENCE FROM N.A.
 RP TISSUE=Skin;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]

RP SEQUENCE FROM N.A.
 RC TISSUE=Skin;
 RX Strausberg R.;

RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC071945; AAH71945.1; -

SQ Nucleocapsid: Ribonucleoprotein.
 SEQUENCE 320 AA; 34180 MW; 9069C6B408DE1AF3 CRC64;

Query Match 93.8%; Score 228; DB 2; Length 320;

Best Local Similarity 100.0%; Pred. No. 2.2e-16; Mismatches 0; Indels 0; Gaps 0;

Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NOSSNFGPMKGNFGGRSSGPGYGGGQYFAKPRNQGCGG 40

DB 268 NOSSNFGPMKGNFGGRSSGPGYGGGQYFAKPRNQGCGG 307

RESULT 13

ROAL HUMAN STANDARD; PRT; 371 AA.
 ID ROAL HUMAN STANDARD; PRT; 371 AA.
 AC P09651.

DT 01-MAR-1989 (Rel. 10, Created)

DT 01-AUG-1990 (Rel. 15, Last sequence update)

DT 01-OCT-2004 (Rel. 45, Last annotation update)

DE Heterogeneous nuclear ribonucleoprotein A1 (Helix-stabilizing

protein) (Single-strand binding protein) (hnRNP core protein A1).

GN Name=HNRPA1;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN SEQUENCE FROM N.A. (ISOFORM A1-A).

RC TISSUE=Liver;

RX MEDLINE=89342435; PubMed=2760922;

RA Biamonti G., Buvoili M., Bassi M.T., Morandi C., Cobiainchi F., Riva S.;

RT "Isolation of an active gene encoding human hnRNP protein A1. Evidence

for alternative splicing.";

RL J. Mol. Biol. 207:491-503(1989).

RN [2]

RP SEQUENCE FROM N.A. (ISOFORM A1-A).

RC TISSUE=Fibroblast;

RX MEDLINE=88233978; PubMed=2836799;

RA Buvoili M., Biamonti G., Ghetti A., Riva S., Bassi M.T., Morandi C.;

RT "cDNA cloning of human hnRNP protein A1 reveals the existence of

multiple mRNA isoforms.";

RL Nucleic Acids Res. 16:3751-3770(1988).

RN [3]


```

FT TURN 73 73
FT STRAND 77 78
FT TURN 79 80
FT STRAND 81 82
FT STRAND 84 87
FT STRAND 105 112
FT TURN 113 114
FT TURN 117 124
FT TURN 125 127
FT STRAND 130 137
FT TURN 139 141
FT STRAND 144 152
FT TURN 155 162
FT TURN 163 163
FT STRAND 168 169
FT TURN 170 171
FT STRAND 172 173
FT STRAND 175 178
SQ SEQUENCE 371 AA; 38715 MW; B3EEFA5AE1DB7C26 CRC64;

Query Match
Best Local Similarity 93.8%; Score 228; DB 1; Length 371;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NOSSNFGPMKGNFGRRSSGPGYGGGQYFAKPRNQGYYGG 40
DB 319 NOSSNFGPMKGNFGRRSSGPGYGGGQYFAKPRNQGYYGG 358

RESULT 14
ROAL MACMU STANDARD; PRT; 319 AA.
AC Q28521;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Heterogeneous nuclear ribonucleoprotein A1 (Helix-destabilizing
DE protein) (Single-strand binding protein) (hnRNP core protein A1).
GN Name=HNRPA1;
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
GX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Tracheobronchial epithelium;
RX MEDLINE=93192320; PubMed=8448206;
RA An G., Wu R.;
RT "cDNA cloning of a hnRNP A1 isoform and its regulation by retinol in
RT monkey tracheobronchial epithelial cells.";
RL Biochim. Biophys. Acta 1172:292-300(1993)
CC -1- FUNCTION: Involved in the packaging of pre-mRNA into hnRNP
CC particles, transport of poly(A) mRNA from the nucleus to the
CC cytoplasm and may modulate splice site selection.
CC -1- SUBCELLULAR LOCATION: Nuclear. Shuttles continuously between the
CC nucleus and the cytoplasm along with mRNA. Component of
CC ribonucleosomes.
CC -1- SIMILARITY: Contains 2 RNA recognition motif (RRM) domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; M84334; AAB01436.1; -.
CC FIR; S30192; S30192.
CC HSSP; P09651; 2UPI.
CC InterPro; IPR000504; RNA_rec_mot.
CC Pfam; PF00076; RRM_1; 2.

```

```

DR PROSITE; PS50102; RRM; 2.
DR PROSITE; PS00030; RRM_RNP_1; 2.
KW Methylation; mRNA transport; Nuclear protein; Phosphorylation; Repeat;
KW Ribonucleoprotein; RNA-binding; Transport.
FT INIT MET 0 0
FT DOMAIN 3 93
FT DOMAIN 94 184
FT DOMAIN 13 96
FT DOMAIN 104 183
FT DOMAIN 217 239
FT DOMAIN 194 319
FT DOMAIN 267 304
FT MOD_RES 5 5
FT MOD_RES 193 193
FT MOD_RES 312 312
FT MOD_RES 307 312
FT POLY-SER.
SQ SEQUENCE 319 AA; 34089 MW; 5A04D9E3BFD969E3 CRC64;

Query Match
Best Local Similarity 91.4%; Score 222; DB 1; Length 319;
Matches 39; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NOSSNFGPMKGNFGRRSSGPGYGGGQYFAKPRNQGYYGG 40
DB 267 NOSSNFGPMKGNFGRRSSGPGYGGGQYFAKPRNQGYYGG 306

RESULT 15
Q9H4S1
ID Q9H4S1 PRELIMINARY; PRT; 70 AA.
AC Q9H4S1;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE BA438F9.2 (Novel protein similar to heterogeneous nuclear
DE ribonucleoprotein A1 (HNRPA1)) (Fragment).
GN Name=BA438F9.2;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Bates K.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL390738; CAC12722.1; -.
DR GO; GO:0030529; C:ribonucleoprotein complex; IEA.
DR GO; GO:0019013; C:viral nucleocapsid; IEA.
KW Nucleocapsid; Ribonucleoprotein.
FT NON_TER 1 1
SQ SEQUENCE 70 AA; 7084 MW; 44379598BA26329E CRC64;

Query Match
Best Local Similarity 90.1%; Score 219; DB 2; Length 70;
Matches 39; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NOSSNFGPMKGNFGRRSSGPGYGGGQYFAKPRNQGYYGG 40
DB 18 NOSSNFGPMKGNFGRRSSGPGYGGGQYFAKPRNQGYYGG 57

Search completed: October 22, 2004, 13:46:24
Job time : 197 secs

```